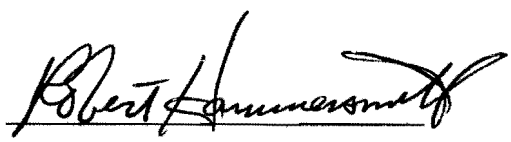
Oxytricha Cyst cDNA Library Analysis and RNAi Examination

An Honors Thesis (HONRS 499)

By

Daniel P. Kane

Thesis Advisor
Dr. Robert L. Hammersmith



Ball State University
Muncie, Indiana

May 2006

May 6, 2006

Acknowledgements

Research is hardly ever solely one individual's work, but relies on building on the work of others and asking for advice along the way. First, I want to thank Patrick Edger for his previous work in isolating mRNA from resting cysts for use in my project. I also want to thank Dr. Iain Cartwright for exposing me to RNAi research and inspiring me to try to translate that work into a new organism, *Oxytricha*. Thirdly, I would like to thank all those in the Biology department who were willing to lend an ear and offer advice on approaching different aspects of this thesis.

Special thanks is needed for my advisor, Dr. Hammersmith, who made himself so approachable that it only took a tiny inquiry as a freshman on how to get started in research that led me into his lab for three years of research, intellectual and personal growth. He was especially helpful while working on this project, helping me through each obstacle that is common in scientific research.

Finally, thanks to those who provided me with a break during the project for growth in other areas. My second family, Bob, Ann and Pat, provided great comic relief when needed. Also thanks to my family, Mom and Dad, and to all my friends who helped me get to where I am today.

Abstract

Oxytricha trifallax is a fascinating organism for genetic study. This ciliate undergoes precise disassembly and reassembly of ciliary structures during several different developmental processes by mechanisms not well defined or understood. Understanding genetic mechanisms behind intracellular reorganization is possible by creating cDNA libraries and finding cDNAs specific to certain stages of development. It is the goal of the laboratory to create these cDNA libraries for multiple developmental processes and stages. But before that can occur, the protocol needs to be analyzed for its feasibility and all troubleshooting worked out. It is one aim of this thesis to analyze the data from a cyst cDNA library to determine what may need to be altered for future library construction. The second aim of this thesis is to determine the feasibility of another method for analyzing genes important to specific stages of development. With the increased interest in RNA interference, this thesis explores the possibility of using dsRNA at the small scale, single-cell level to interfere with gene expression for purposes of analyzing genes found to be important to developmental stages through library creation.

Introduction

Oxytricha trifallax is a small ciliated protozoan that has been utilized as a model system in a number of different research areas. One of these areas is trying to understand the cellular processes involved in the patterning of surface structures. Four major developmental processes (division, reorganization, conjugation and cystment) involve major disassembly and re-assembly of surface ciliary structures that then become arranged in a relatively precise global pattern characteristic of the species (1). New developing structures have a spatial/ temporal relationship to one another and to the whole (global patterning). Previous research has demonstrated that this overall global pattern depends upon a combination of interactions of developmentally regulated genes and a non-genic inheritance system referred to as cytotaxis - the role of pre-formed structure in the organization of new structure (2).

The process of encystment, which produces a dormant, environmentally resistant stage termed a cyst, is especially unique and involves a complete degeneration of all ciliary structures and visible surface pattern, including all microtubules (3). This presents questions as to what genetic mechanisms are involved in the proper rearrangement of ciliary structures that occurs during excystment (the return to its vegetative state). Finding major genes that influence the regulation, interaction and expression of these cellular patterning pathways are of major theoretical importance.

Dr. Robert Hammersmith has been working on the long-term project in *Oxytricha* that includes the goal to isolate developmentally regulated genes that influence these various developmental programs involved in intracellular patterning and organization. One way to identify developmentally important genes is to look for unique messenger RNAs (products of genes) that are expressed only during a specific developmental stage. This has led to the lab's current project of developing procedures for the creation of expression libraries in *Oxytricha* from various developmental stages.

Expression libraries are a collection of DNA clones (cDNA) that are derived from messenger RNAs being produced by cells of a certain type or at a certain developmental stage. Simplistically, this involves the isolation of messenger RNA from specific staged cells, then the production of copy DNA by reverse transcription, and the introduction of that cDNA into large DNA cloning vectors. A comparison of expression libraries from different staged cells can reveal genes that are expressed only within a specific developmental stage.

Commercially available kits (Invitrogen) now simplify this process provided that sufficient quantities of stage-specific cells are available. mRNAs can be isolated from ciliates in sufficient quantities to produce cDNA libraries (4). These initial experiments utilized cysts, since the cyst stage represents a unique and easily induced developmental stage in which relatively large quantities of staged material can be obtained. For many of the unique developmental stages, however, large quantities of synchronized, stage-specific material (various time points in the encystment or excystment stages) are difficult to obtain. This is due to the variation from cell lines and from cell to cell in how

quickly these processes occur, making it difficult to get mass groups at a specific stage. Consequently, the development of new protocols utilizing small quantities of synchronized material for the isolation of mRNAs and their use in construction of expression libraries needs to be undertaken. Future research will involve working out the protocols for the isolation of the messenger RNA from limited quantities of materials and their use with the expression library kits. While theoretically it should be possible to develop procedures that can use a single cell of *Oxytricha* (5), initial experiments will likely attempt to develop procedures that use several hundred cells that can be easily synchronized and isolated by hand using micropipettes.

While preparations are underway to obtain libraries from all stages of *Oxytricha* to determine developmentally-regulated genes, the first part of this thesis is to optimize the protocol for creating a library using the large-scale approach and to analyze select clones. mRNA has been isolated from *Oxytricha* during the cyst stage previously in the lab. This mRNA was used for the construction of a cDNA library, where select clones were sent out for sequencing and then analyzed.

With the successful creation of libraries for each stage of *Oxytricha* development, a direction for further research is determining the function of genes found to be specific to each developmental stage. This comprises the second part of this thesis, to see if an RNA interference (RNAi) approach could be taken to down-regulate expression of genes determined to be developmentally important to understand their precise function. To understand how to attack mRNA expression *in vivo*, a further review of this ciliate is needed.

Oxytricha possess two types of nuclei: micronuclei, which are the germinal nuclei, and function in sexual reproduction; and macronuclei that serve as somatic nuclei regulating and coding for the messenger RNAs (and hence the proteins) expressed during all developmental stages. Micronuclei undergo meiosis to produce gametic nuclei that then participate in a reciprocal fertilization during conjugation resulting in a new diploid zygote nucleus. This new zygotic nucleus divides mitotically to produce two nuclei; one that will become the new micronucleus and hence set aside as the germinal reserve, and the other that undergoes a complex series of alterations to become the new somatic macronucleus, as the old macronucleus degenerates (6,7). The alterations associated with the production of the new macronucleus is a complex developmental program, involving amplifications of DNA and elimination of approximately 90 % of the micronuclear DNA sequences (Internal Eliminated Sequences) (8). Rearrangements of the remaining DNA sequences then occur to form small gene-size mini-chromosomes that are then re-amplified to form a highly polyploid macronucleus. The DNA rearrangements involve poorly understood mechanisms but involve complex interactions in which various sequences are pieced together, sometimes with reverse orientation, to form the functional gene. Additionally, in some cases, a single micronuclear gene can be demonstrated to produce multiple types of different macronuclear genes by alternate DNA rearrangements (7).

In 1998, Andrew Fire and Craig Mello, along with their colleagues, injected double stranded RNA (dsRNA) into worms and, much to their surprise, the double stranded RNA inhibited the genes that generated the RNA in the first place. This is the result of the enzyme named Dicer that recognizes dsRNA and cuts it into smaller RNA molecules (9). These small RNA molecules result in the RNA-mediated interference (RNAi) of gene expression by triggering the degradation of the homologous RNAs in a cell or organism. This has been demonstrated in the ciliate *Paramecium* by feeding them dsRNA-expressing bacteria (10). Follow up experiments also demonstrated that it was possible to introduce dsRNA-expressing bacteria to *Oxytricha* (11). However, in these experiments, the ciliates had to become adapted to *E. coli* as a food source to uptake the dsRNA.

Therefore, instead of trying to get *Oxytricha* onto a bacterial diet, the second half of this thesis examines whether one could grow individual ciliates in media under a droplet of oil (to prevent evaporation) and introduce dsRNA directly with the normal algae food source. If the method worked, then the data obtained from libraries created by our lab and from previous research could be tested by this method. Previous research of mRNA in cysts reveal that the transcripts available before excystment include more than structural or housekeeping genes but also unique mRNAs that encode proteins involved in signaling pathways (12). As a cyst, the ciliate is relatively inactive, therefore one could assume transcripts present that initiate signaling would be excystment specific and could have been transcribed during encystment. Sequence analysis of transcripts from cysts readying for excystment reveal one transcript to be a homologue of CROC-1 (4). To determine the function of these transcripts and others deemed important by library construction in the lab, RNAi could be utilized - if the method proves successful. To test the method, dsRNA can be designed to a commonly expressed gene (actin) and determine if function is lost (phenotypic effect on ciliary and cytoskeletonal formation could be detected). It is important to use sequence from the macronucleus and not the micronucleus when designing primers to create the dsRNA, since, as shown previously, it is the macronucleus that is actually being expressed.

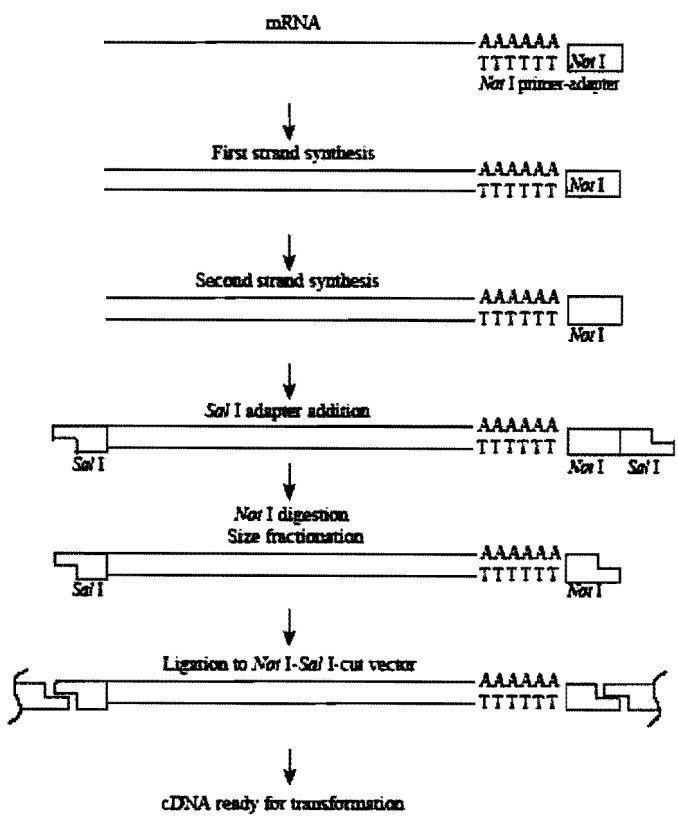
In summary, this thesis has two purposes. The first is to create a cDNA library for *Oxytricha* using mRNA from the cyst stage and sequence and analyze select clones to find genes important to this specific developmental stage. Secondly, this thesis aims to show whether RNAi can be induced in *Oxytricha* by direct feeding of dsRNA in hopes of finding a method to test the function of genes found specific to each developmental stage.

Materials and Methods

cDNA Library Analysis

Total RNA was previously isolated from $3x10^5$ JRB 310 *Oxytricha* cells in the cyst stage with Trizol reagent (Life Technologies, Cat #15596). Poly A^+ RNA was then selected using the Messagemaker reagent assembly (Life Technologies, Cat #10298-026), according to the manufacturer's instructions. The Poly A tail is a modification that is unique to mRNA and selecting for RNA with a Poly A tail removes rRNA and tRNA contaminants. Final mRNA was stored in glycogen at -20°C.

To not lose any starting mRNA sample, and considering the efficiency of the Trizol extraction, the entire mRNA sample was used for the cDNA library construction. It was estimated that the sample contained about 5 µg of mRNA. The mRNA was spun, supernatant removed, and pelleted mRNA resuspended in 5 µl of DEPC-treated water. This entire volume was used for cDNA construction according to manufacturer's instructions using Superscript™ Plasmid System with GATEWAY™ Technology for cDNA Synthesis and Cloning (Invitrogen, Cat #18248). Any solutions that were not provided were made according to the manufacturer's specifications or according to a lab-bench reference book (13). Utilizing mRNA's poly A tail, primers can be designed for opposite strand synthesis. Addition of Not I and Sal I adapters allow cDNAs to be ligated into a vector for transformation (see diagram, cDNA diagrams from Invitrogen manual). ^{32}P was not used to quantify strand synthesis at any step.



cDNA was ligated into vector pCMV•SPORT 6 and transformed into DH5α cells (vector is taken in by bacteria to be replicated to make many copies). Transformation was performed using the combined protocol of Section 3.9 of the Superscript™ Plasmid System and instructions included with the MAX Efficiency® DH5α™ Competent Cells (Invitrogen, Cat #18258-012). The cDNA library was expanded using Protocol 5.1 of the Superscript™ Plasmid System, according to manufacturer's instructions. Successful transformation was screened by ampicillin resistance (coded for by the pCMV•SPORT 6 vector) before and after expansion.

Twelve individual colonies (samples bsuCST-1 through bsuCST-12) were grown up in 5 ml of LB broth with Ampicillin for 16 hours. 2 ml of the growth was used for plasmid isolation. The remaining aliquots, as well as all other individual clones and transformation aliquots were stored in 2x LB with glycerol (12.5%) at -80°C. The twelve clone growths had plasmid DNA isolated using QIAprep® Spin Miniprep Kit (Qiagen, Cat #27104). Final plasmid DNA was eluted off column using 50 µl of Qiagen's EB Buffer (10mM Tris-Cl, pH 8.5).

A 1:1000 dilution was taken of each plasmid sample to calculate concentration in a spectrophotometer. Entire samples were then shipped to The Ohio State University Plant-Microbe Genomics Facility for sequencing. Sequence data was then analyzed using BLAST searches through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).

RNAi Examination

The *Oxytricha* food source, *Chlamydomonas allensworthii* (Cat #2717), was purchased from The Culture Collection of Algae at the University of Texas at Austin. The algae was grown in an MES-volvox medium as described by the University of Texas. JRB 310 *Oxytricha* cells were expanded by feeding *Chlamydomonas* in a 2 L culture flask.

The NCBI website was utilized to obtain macronuclear sequence for the actin gene (*Oxytricha trifallax* B macronuclear actin I gene, complete sequence; U63568). Primers were designed with the aid of FastPCR. Primers were further analyzed using IDT OligoAnalyzer 3.0 (Integrated DNA Technologies) and purchased through Integrated DNA Technologies. The aim of the analysis was to look at primer sets with close melting temperatures (T_m) and little secondary structure that would interfere with the PCR reaction.

PCR was run on JRB 310 *Oxytricha* DNA dilution samples according to the 50 µl reaction described by DuBois, and Prescott, 1997 (14), using the lowest annealing temperature. Primer, dNTP and $MgCl_2$ concentrations, as well as annealing temperatures were varied to optimize the PCR reaction.

When PCR is finally optimized, PCR samples can be purified using a High Pure PCR Product Purification Kit (Roche, Cat #11732668001), and a second round of PCR run using primers with T7 promoter regions (T7 promoters are used for dsRNA synthesis). dsRNA could then be made using the final purified PCR product using the MEGAscript® RNAi Kit: High Yield Transcription Kit for dsRNA Preparation (Ambion, Cat #1626), according to manufacturer's instructions.

An experiment could be set up with individual *Oxytricha* cells in four experimental groups. Cells would be grown under a drop of vegetable oil with the same amount of *Chlamydomonas* food source in each group. Group 1 would be grown as a control, with no dsRNA introduced. Group 2 would be grown with dsRNA to actin introduced. Group 3 would serve as a negative control, dsRNA created to gene not found in *Oxytricha*, thus should have no effect (e.g. dsRNA designed against bovine hemoglobin). Group 4 would be a positive control, as a chemical (cytochalasin B) would be introduced that is known to affect *Oxytricha* shape and movement (even leading to cell death). Swimming patterns of *Oxytricha* would then be observed to see if actin expression was affected and RNAi could be induced by straight feeding of dsRNA, and not relying on an E.coli intermediate.

Results

cDNA Library Analysis

Before expansion of the cDNA Plasmid Library, dilutions of the transformed DH5α cells were plated on agar plates with and without ampicillin selection to determine transformation efficiency.

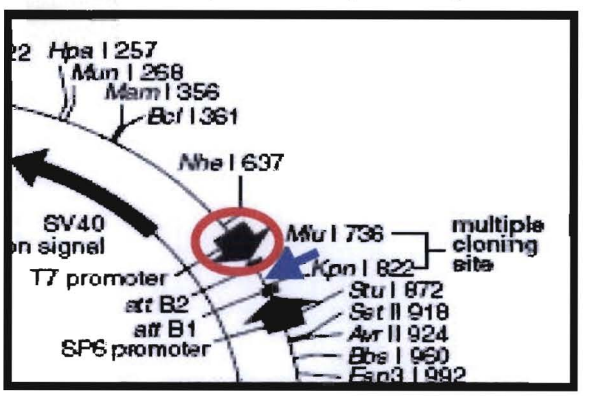
Agar Plates, No Ampicillin		Agar Plates, Ampicillin	
10^{-4} dilution	Lawn	10^{0} dilution	Lawn
10^{-4} dilution	Lawn	10^{0} dilution	Lawn
10^{-6} dilution	286 colonies	10^{0} dilution	Lawn
10^{-6} dilution	297 colonies	10^{-1} dilution	127 colonies
		10^{-1} dilution	113 colonies
		10^{-1} dilution	102 colonies

Without Ampicillin: Average 291.5 colonies at 10^{-6}, therefore, ~ 291 million colonies on an undiluted plate.

With Ampicillin: Average 114 colonies at 10^{-1}, therefore, ~ 1,140 colonies on an undiluted plate.

This gives a transformation efficiency of 1,140/291,000,000 or 0.0004%. This means that out of every million DH5α cells only 4 DH5α cells took up the pCMV•SPORT 6 vector. However, using ampicillin selection allows only cells with the vector to survive to be screened for cDNA inserts.

After the selected twelve individual colonies were grown overnight and had plasmid DNA isolated using Qiagen's Miniprep Kit, spectrophotometer readings were taken. 1:1000 dilutions were necessary to reserve eluted DNA for sequencing, but makes accurate readings difficult. An average of all 12 spectrophotometer readings was taken, since all 12 colonies underwent the same growth conditions, so would be expected to have similar final plasmid concentrations. The remaining plasmid DNA isolation was shipped to The Ohio State University for sequencing. The concentration sent was 300 ng/μl. Based on the directionality of the library system, using the T7 promoter primer (red circle) would give the sequence of the cDNA inserts (blue arrow) starting from the 5' end on down towards the 3' poly A tail.

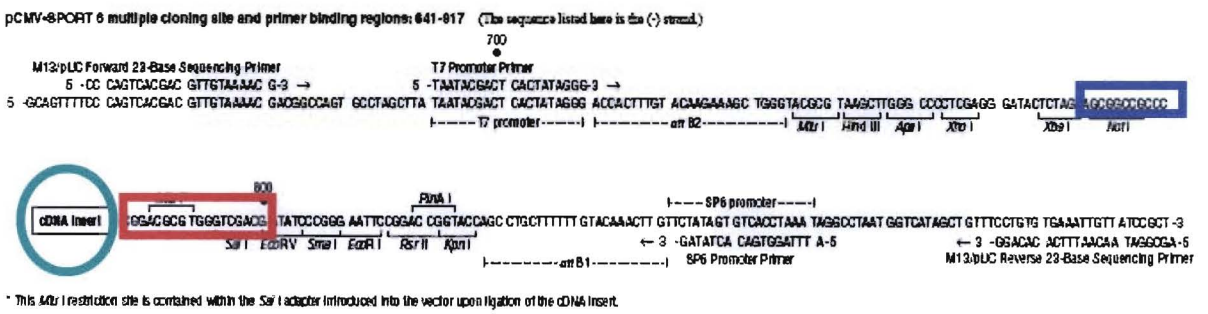


Sequencing results came as text, each base pair listed, and as a chromatogram. Eleven out of the twelve samples sent in for sequencing came back with good sequencing. The following is an example of the sequencing text received for all samples (from sample bsuCST-11).

bsuCST-11:

NNNNNNNNCTGGGTACGCGTANCTTGGGCCCCTCGAGGGANACTCTAGAGCGGCCGCCCTTTTTTTTTTT
TTTTATGCTTTAATGATTGATAGAAAAATTTAAAGATTAAAGATCAAAATTTATCAAAGATTATATCTAG
AATGGTGTCTAAACATAAAATTANAATCNCTACTTTCTCTTGAATCTTATGTGATTCTGTATAGTCCATCA
TCTGATTGTTAATAACATCATTCNCAATCCATGATACTTCACAATTTTGATATAGTCNAACAANAGCTATG
CCANANTAGTGAATGAAATTAATATAAAGCTTGGTTTCACTCTCAATTACAAAAGCGTAATANACTCTCA
AGAATTTGTTCAAGTATTAGTTTCNTCTTCAATATAGGCTGTAGTGANAATTTTCTGNGNGTTATGTTCAT
GCGTATGATTTTTCTTTAATATCTCAATATTGTTTATGCTGTTTAAATTGTGAGTATTATTGAATTTTGAAT
TTCTATTCTTTGTCAATAAATCAAAAATCTATTGCTTAAGATTTAANAAATGTTCTTCGGACGCGTGGGTC
GACGATATCCCGGGAATTCCGGACCGGTACCAGCCTGCTTTTTTGTACAAACTTGTTCTATAGTGTCACCT
AAATAGGCCTAATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGTCCCGCGGCCTAGGCTAGAGTCC
GGAGGCTGGATCGGTCCCGGTGTCTTCTATGGAGGTCAAAACAGCGTGGATGGCGTCTCCAGGCGATCTG
ACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAAT
GGGGCGGAGTTGTTACNACATTTTGGAAAGTCCCGTTGATTTTGGTGNCAAACAACTCCCATTGACGTCA
ATGGGGTGGANACTTGGAAATCCCCGTGAGTCAACCGCTATCCNCGCCCATTGATGTACTGCCAAACCGC
ATCNCCATGGNAANAGCGATGACTAATNCGTAGATGTACTGCCNAGTANNAAGTCCCATNNNCNTGTAC
TGGGNNNANGNCNGNNGNNATTTNNCGTCATTGACGTCATNGGGGCGTANTNGNNNNNGANNNNNNTG
NTNNACTNNNNTGGCNNTTNACCGTANNACTNNNCCCNNNNNGTNANNNNNGTCNNNTNNNTTNNNNNN
NNNANNNCNTATTGANNTCANGNNGGGGNCNNNGNGNNNNNAGNNNGNNNNNCGNANTNNNNNCNNN
NNNNNNNNGNNNNNGCCTNAANNCNNTTTNNGNNATNNNNNNANNNNNNNANNNNNNN

Within the instructional manual for Invitrogen's Superscrpt™ Plasmid System is the sequence for the pCMV•SPORT 6 vector to determine where in that above sequence is the actual sequence for the cDNA insert. The area highlighted in blue is the upstream Not I cut region and the area highlighted in red is the Sal I cut region. The area highlighted in green is where the cDNA insert is.



* This Mlu I restriction site is contained within the Sal I adapter introduced into the vector upon ligation of the cDNA insert.

Knowing the actual sequence of our vector, we can now go further into the rough sequencing results and determine where the cDNA insert lies. Using the same color coding, the sequence highlighted in blue is the 5' upstream region of the vector and the red is the 3' downstream region of the vector. Realizing this, it was found that only 3 of the remaining eleven sequences had a cDNA insert in the vector. The underlined region of the example is what can then be used for BLAST searches and the upstream and downstream regions deleted.

bsuCST-11:

GCGGCCGCCCTTTTTTTTTTTTTTTATGCTTTAATGATTGATAGAAAAATTTAAAGATTAAAGA
TCAAAATTTATCAAAGATTATATCTAGAATGGTGTCTAAACATAAAATTANAATCNCTACTTT
CTCTTGAATCTTATGTGATTCTGTATAGTCCATCATCTGATTGTTAATAACATCATTCNCAATC
CATGATACTTCACAATTTTGATATAGTCNAACAANAGCTATGCCANANTAGTGAATGAAATTA
ATATAAAGCTTGGTTTCACTCTCAATTACAAAAGCGTAATANACTCTCAAGAATTTGTTCAAG
TATTAGTTTCNTCTTCAATATAGGCTGTAGTGANAATTTTCTGNGNGTTATGTTCATGCGTATG
ATTTTTCTTTAATATCTCAATATTGTTTATGCTGTTTAAATTGTGAGTATTATTGAATTTTGAAT
TTCTATTCTTTGTCAATAAATCAAAAATCTATTGCTTAAGATTTAANAAATGTTCTTCGGACGC
GTGGGTCGAC

A Blastx search (with the ciliate genetic code option selected, as *Oxytricha* reads two of the three typical stop codons as the amino acid glutamine) was performed on all three of the sequences determined to have a cDNA insert. A Blastx search is a translated query vs. the protein database search. This search revealed no matches, matches to theoretical proteins and other matches with high E values, with low similarity between sequences. The following are the results obtained in the Blastx searches.

bsuCST-6:

Sequence:

NCKYKWTNTNTNNNTNNTNNNNNTTNTTGNTNCNACNNTTNCCNGGAANNTCNGAACNNNTNCCNNCNT
NTTTTNTTTTNTTTNTNGNCCAANANNNNCNNCNAAANAGNCNNANGGNCCNNACCNGTTNCCGGNGNN
AATTNGTNTNCCNNCCCGGGNCNAAGGNTAAANCCNGAAGGTTGNATCGNCCCGGGGNTCTNCNANGGN
GTNNAAAANGGNGGGNAGNGNNNCNCCNNNNAANTTNACGTTTNANNAAANAANTCCGNTTAANANNN
CCNTCNNNNNNAACCNCCCNNCNNCCNTTTTNNNNNNNNGGNNNGAANTTNTNNNAANNTTTGGAAANN
CCNNTTNNATTTNGNNGCCAAANNAANNNCCCTNNGNNNCNNNGGGGGGGAANANTGGNAANCCCCCN
NNANNNAACCCGNNNNCCNNCCCNNTNNNNNNNCNCCNAAANCNNNTTCCCNAGGNNAATANCNATGA
CNATAACGAANANTNNNNGCNANNNANAAAGNTCCNNAAGGGCCTTGTATNGGGNNNANGGCNGGGNG
GNCNNTTANCCGTNNTNNNCNCCAANAGGGGGNTNACTTGNNNNAGAANANGCTNGATGTANGGCNNG
NGGNNCNGTTNACNGNAANNCNCCCCCCNNTTGNCNCANNGGAANANCNCNAANTGNNGTNNCTANGG
NCNCTNNNNTNNTNATNGNCNCCATGNGGGGGGGNTNGTGGGGGNGGNCCNCCNGGGNGGNCNNTCNN
CCNGNANTTTNTTNNANCNNCTTGNCNTNAANNGNNGGGAAGGANCNCCNNAAAANNNCCCTNTTTTTN
ANGGNTNNTGGNCNNGNNNAAAANGGNNNANNAAAANGNNNGGNGGACCNTTTCCGCGGGNNANNGG
GGNGNGNACCCCNNNTNGATTTNTTTTTTNNANNTNNNTTCAAANGGGNNTCCNANNNGNNNNCNNNAN
CTGGNTANNNCNTANAAAATTNTAAAANANNNNNN

Blastx results:

Query=
Length=992

No significant similarity found. For reasons why, **click here**.

bsuCST-10:

Sequence:

NCKYKNTYTWNNNYTNNTCNNNNNNNNTGNNNNNANNANTTCCNNGNANTTCNGGNNNNNANCCNNNN
TNTTTTTTTGNNGNNNTTGTNCNNNNNNGTCNCCTAAATNGNCCNANTNNTNNTANNNGTGNNNNGNNN
NAAATNNTNANCCANCCCGCCTTTTNGCTANANTCNGAAGNTTGGATCGGNCCCNGAGNNTNNNANTGN
GTCCAAAANGNNNNGNANGNNNNCNCCNATTATTCNGCCGNTCNATTANACNANNCCTGTTTNTANANA
CNTCTTTTNNACNNCCCANCNNCCANTTTGNCTCCNNGNNTTTNANTTGTNNNNANATTTNGNAAAANCT
TTTNGATNNNNNNGCNNAANTTNNCNCNANNTGNNGTTNNNGNNGNGNACANTTGNNNNNCCNCNTNN
NNNANACCGNNANCCANNNNAATNNCTTNNCNGNNNAANNTNGNNCNCAANGNAAAANNGAATNACTA
NNCCGTNNANTNANNGGGAANTATGANGNNCCNATANNNNTNGNATTGGGANTTGTGCCGGGCANNCC
NTTNANCGNCNTTGANNCCGAANGGGCCNCACTTGGNNTNNNTCNACTNNATGTTNTGCCNGNTGNACN
GNTNACCGNANAAANTCCNGCCANTGNNNCNAANGANNANTCCTAANTGNNNNNACTATGGNANCNAN
NNNNNTNATTNACTNCANTGTGGGNGNGNNNGNNCGNGCCTAGCNAAGCGGNCNNNTNNNGGNANNTN
NTGNNGNNNCTNGCNTCNANNNCTTGAANGGNCCCNNNAANNCGNCCTNGTTNAAAGNTNANTGTCNTG
GNAANANGGNTTACNNTNGANGNANNTGGAAATNTC

Blastx results:

Query=
Length=859

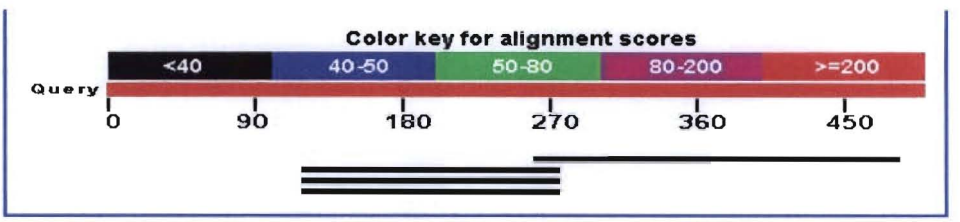
No significant similarity found. For reasons why, click here.

bsuCST-11:

Sequence:

CCTTTTTTTTTTTTTTTATGCTTTAATGATTGATAGAAAAATTTAAAGATTAAAGATCAAAATTTATCAAA
GATTATATCTAGAATGGTGTCTAAACATAAAATTANAATCNCTACTTTCTCTTGAATCTTATGTGATTCTG
TATAGTCCATCATCTGATTGTTAATAACATCATTCNCAATCCATGATACTTCACAATTTTGATATAGTCNA
ACAANAGCTATGCCANANTAGTGAATGAAATTAATATAAAGCTTGGTTTCACTCTCAATTACAAAAGCGT
AATANACTCTCAAGAATTTGTTCAAGTATTAGTTTCNTCTTCAATATAGGCTGTAGTGANAATTTTCTGNG
NGTTATGTTCATGCGTATGATTTTTCTTTAATATCTCAATATTGTTTATGCTGTTTAAATTGTGAGTATTAT
TGAATTTTGAATTTCTATTCTTTGTCAATAAATCAAAAATCTATTGCTTAAGATTTAANAAATGTTCTT

Blastx results:



Sequences producing significant alignments:		Score (Bits)	E Value
gi|82500844|ref|ZP_00886238.1|	ATPase, E1-E2 type [Caldicellu...	33.5	3.4
gi|46906569|ref|YP_012958.1|	cell wall surface anchor family ...	32.0	9.8
gi|47097218|ref|ZP_00234781.1|	internalin protein lin1204 [Li...	32.0	9.8
gi|47091467|ref|ZP_00229264.1|	cell wall surface anchor famil...	32.0	9.8

To increase the E values to better determine what the proteins/functions of the cDNA inserts are and how they might relate to the cyst stage of ciliates, the chromatograms of the three sequences were used to fill in the N (unknown) values of the sequence data. Chromatograms are a visual presentation of the sequence, base by base, and when peaks were too close to call, the tallest peaked base was selected to replace the N value. With this performed, the new sequences to perform Blast searches were as follows.

bsuCST-6:

Sequence:

CCTTTATNTNTATTTNTTCNNNNTTNTTGNTNCNACAATTTCCNGGAANTTCNGAACGGTTACCNNCNTGT
TTTTTTTTNTTTTTNGNCCAAAANGGNCACCTAAANAGGCNNATGGNCCNNACCNGTTNCCGGGGNNAAT
TNGTNTNCCNNCCCGGGNCNAAGGNTAAAGCCNGAAGGTTGNATCGGCCCGGGGGTCTTCNANGGAGTC
CAAAANGGNGGGAAGGGCGCCNCCNNNGAATTTGACGTTTCATNAAACAANTCCGGTTAAAATAACCNT
CCNNCNNAACCCCCCANCNNCCATTTTGANCCANGGGGTGGAANTTNTAACAANATTTGGAAANACCCT
TTNNATTTTGGNGCCAAANNAANNNCCCTTNGANGCNANGGGGGGGAANATTGGNAANCCCCCNNGAN
NNAACCCGNNANCCACCCCNNTTGNNNNNCGCCNAAANCNNCTTCCCNAGGNNAATANCNATGACNAT
AACGAANANTNANNGCNAANNANAAAGNTCCNNAAGGGCCTTGTATNGGGNNNANGGCNGGGNGGNCN
NTTANCCGTNNTNNNCNCCAANAGGGGGNTNACTTGGAATAGAANANGCTTGATGTANGGCNAGGGGG
NCNGTTNACNGNAANNCNCCCCCCNNTTGCCNCANNGGAANACCNCNAATTGGNGTNNCTANGGNCACT
NNNNTNNTNATNGNCTCCATGGGGGGGGGNTCGTGGGGGGGGNCCNCCNGGGNGGNCNNTCNNCCNGN
ANTTTNTTNNANCNNCTTGNCNTNAANNGNNGGGAAGGANCNCCNNAAAANNNCCCTNTTTTTNANGGN
TNNTGGNCNNGNNNAAAANGGNNNANNAAAANGNNNGGNGGACCNTTTCCGCGGGNNANNGGGGNGN
GNACCCCNNNTNGATTTNTTTTTTNNANNTNNNTTCAAANGGGNNTCCNANNNGNNNNCNNNANCTGGN
TANNNCNTANAAAATTNTAAAANANNNNNN

Blastx:

Query=
Length=992

No significant similarity found. For reasons why, click here.

bsuCST-10:

Sequence:

NCKYKATCTATANYTATTCNNNANNCNTGNNTCGACNATTTCCNGGAANTTCNGGACNGNAA
CCNNCATGTTTTTTTGTNGNNNTTGTNCNNAACTGTCACCTAAATAGGCCNANTGNTNNTANN
NGTGNNNNGNNNNAAATTGTGANCCATCCCGCCTTTTNGCTANAGTCNGAAGGTTGGATCGG
NCCCGGAGTCTNNNATTGNGTCCAAAANGNCGGGCANGNNGNCNCCAATTATTCNGCCGNTC
CATTANACNANNCCTGTTTATANATACNTCTTTTNNACNCCCCANCNNCCANTTTGNCTCCNN
GNTTTTNANTTGTNNNNANATTTGGNAAAANCTTTTNGATTNNGATGCNNAATTTGNCGCCAN
TTGNGGTTTTNGNNGNGNACANTTGNNNNNCCNCNTGAAANAAACCGNGANCCANNNNAAT
NNCTTCCCNGAAAAATCTNGTNCNCAANGTAAAAANGAATGACTANNCCGTNNANTNANNGG
GAATTATGANGNNCCNATANNTNTNGNATTGGGANTTGTGCCGGGCANNCCNTTNANCGNCN
TTGANNCCGAANGGGCCNCACTTGGNNTNNNTCNACTTNATGTTNTGCCNGNTGNACTGNTTA
CCGNANAAANTCCNGCCATTGCNNCNAATGANNATTCCTAANTGNNNNNACTATGGNANCNA
NNNNNNTNATTNACTNCANTGTGGGNGNGNNNGNNCGNGCCTAGCNAAGCGGNCNNNTNNN
GGNANNTNNTGNNGNNNCTNGCNTCNANNNCTTGAANGGNCCCNNNAANNCGNCCTNGTTN
AAAGNTNANTGTCNTGGNAANANGGNTTACNNTNGANGNANNTGGAAATNTC

Blastx:

Query=
Length=859

No significant similarity found. For reasons why, click here.

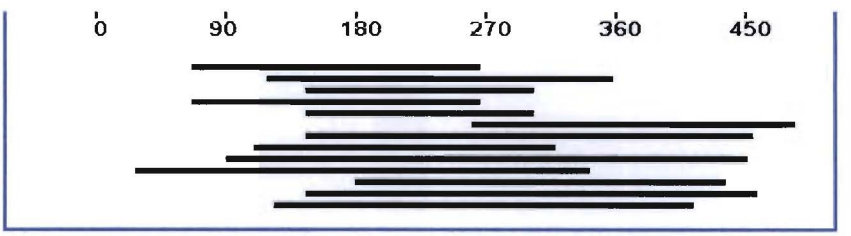
By filling in some of the N bases, there is alignment with the 3' end (in red above) of the vector early on, signifying that there is probably not a cDNA insert in this sample.

bsuCST-11:

Sequence:

CCTTTTTTTTTTTTTTTTATGCTTTAATGATTGATAGAAAAATTTAAAGATTAAAGATCAAAATTTATCAAA
GATTATATCTAGAATGGTGTCTAAACATAAAATTAGAATCTCTACTTTCTCTTGAATCTTATGTGATTCTG
TATAGTCCATCATCTGATTGTTAATAACATCATTCTCAATCCATGATACTTCACAATTTTGATATAGTCGA
ACAAGAGCTATGCCAGAGTAGTGAATGAAATTAATATAAAGCTTGGTTTCACTCTCAATTACAAAAGCGT
AATAGACTCTCAAGAATTTGTTCAAGTATTAGTTTCGTCTTCAATATAGGCTGTAGTGAGAATTTTCTGTG
TGTTATGTTCATGCGTATGATTTTTCTTTAATATCTCAATATTGTTTATGCTGTTTAAATTGTGAGTATTAT
TGAATTTTGAATTTCTATTCTTTGTCAATAAATCAAAAATCTATTGCTTAAGATTTAAGAAATGTTCTT

Blastx:



Sequences producing significant alignments:		Score (Bits)	E Value	
gi|28410441|emb|CAD66827.1|	putative ABC transporter integral...	34.3	2.0	G
gi|23615503|emb|CAD52495.1|	hypothetical protein [Plasmodium ...	33.1	4.4	G
gi|21593009|gb|AAM64958.1|	unknown [Arabidopsis thaliana]	33.1	4.4	
gi|28493105|ref|NP_787266.1|	ABC transporter permease protein...	33.1	4.4	G
gi|15229382|ref|NP_191871.1|	cyclin-dependent protein kinase ...	33.1	4.4	G
gi|82500844|ref|ZP_00886238.1|	ATPase, E1-E2 type [Caldicellu...	33.1	4.4	
gi|83318017|ref|XP_731412.1|	hypothetical protein PY03408 [Pl...	32.7	5.7	G
gi|89302250|gb|EAS00238.1|	oxidoreductase, short chain dehydr...	32.7	5.7	
gi|89299247|gb|EAR97235.1|	hypothetical protein TTHERM_004836...	32.7	5.7	
gi|67975567|ref|XP_668990.1|	hypothetical protein PB300877.00...	32.3	7.5	G
gi|23495112|gb|AAN35443.1|	hypothetical protein PF10_0246 [Pl...	32.0	9.8	G
gi|66812082|ref|XP_640220.1|	hypothetical protein DDB0216408 ...	32.0	9.8	G
gi|82913953|ref|XP_728766.1|	hypothetical protein PY01109 [Pl...	32.0	9.8	G

Of particular interest is the new hit to a ciliate, *Tetrahymena thermophila* (TTHERM). It is a hypothetical protein based off a conceptual translation from a

macronuclear genome sequence with a high E-value (24% identity when looking at proposed amino acid sequence, below), but an interesting find nonetheless.

```
>gi|89299247|gb|EAR97235.1|  hypothetical protein TTHERM_00483670 [Tetrahymena thermophila
SB210]
Length=740

 Score = 32.7 bits (73),  Expect = 5.7
 Identities = 30/122 (24%), Positives = 60/122 (49%), Gaps = 8/122 (6%)
 Frame = +2

Query  92   LNIKLESLLSLESYVILYSPSSDCQQHHSQSMILHNFDIVEQ--ELCQSSE*NQYKAWFH  265
            L + +     LE Y+ L++     Q  ++Q+   +NF I++Q  +L ++ E +Q+K
Sbjct  292  LKVFVNDQQELEDYLELFNKIELTQHIYNQTQKQYNFQILKQNIDLVEAKESSQFK----  347

Query  266  SQLQKRNRLSRICSSISFVFNIGCSENFLCVMFMRMIF-LQYLNIVYAVQIVSIIEF*IS  442
             Q+ K+N    +  S    F I C  N      ++ +F L+   +    Q++S+ E+ I+
Sbjct  348  -QILKKNSGITVPLSNIGQFVIDCVSNIESERLLQFLFALKTQQLEKLNQLLSLEEYKIT  406

Query  443  IL  448
            +L
Sbjct  407  ML  408
```

This was supposed to be a directional library, with cDNA inserting into the vector in only one direction, but it appears that in the sequence of the two remaining samples that there is a poly T (poly A tail) sequence at the beginning. Therefore, analysis was also run on the reverse complements of both samples.

bsuCST-6:

Sequence (Reverse Complement):

TNNTGGNCNNGNNNAAAANGGNNNANNAAAANGNNNGGNGGACCNTTTCCGCGGGNNANNGGGGNGN
GNACCCCNNNTNGATTTNTTTTTTNNANNTNNNTTCAAANGGGNNTCCNANNNGNNNNCNNNANCTGGN
TANNNCNTANAAAATTNTAAAANANNNNNNNCCNTNAAAAANAGGGNNNTTTTNNGGNGNTCCTTCCCN
NCNNTTNANGNCAAGNNGNTNNAANAAANTNCNGGNNGANNGNCCNCCCNGGNGGNCCCCCCCCACGA
NCCCCCCCCCCATGGAGNCNATNANNANNNNAGTGNCCNTAGNNACNCCAATTNGNGGTNTTCCNNTGNG
GCAANNGGGGGGNGNNTTNCNGTNAACNGNCCCCCTNGCCNTACATCAAGCNTNTTCTATTCCAAGTNA
NCCCCCTNTTGGNGNNNANNACGGNTAANNGNCCNCCCNGCCNTNNNCCCNATACAAGGCCCTTNNGGA
NCTTTNTNNTTNGCNNTNANTNTTCGTTATNGTCATNGNTATTNNCCTNGGGAAGNNGNTTTNGGCGNNN
NNCAANNGGGGTGGNTNNCGGGTTNNNTCNNGGGGGNTTNCCAATNTTCCCCCCCNTNGCNTCNAAGGG
NNNTTNNTTTGGCNCCAAAATNNAAAGGGTNTTTCCAAATNTTGTTANAANTTCCACCCCNTGGNTCAAA
ATGGNNGNTGGGGGGTTNNGNNGGANGGTTATTTTAACCGGANTTGTTTNATGAAACGTCAAATTCNNN
GGNGGCGCCCTTCCCNCCNTTTTGGACTCCNTNGAAGACCCCCGGGCCGATNCAACCTTCNGGCTTTANC
CTTNGNCCCGGGNNGGNANACNAATTNNCCCCGGNAACNGGTNNGGNCCATNNGCCTNTTTAGGTGNCC
NTTTTGGNCNAAAAANAAAAAAAAACANGNNGGTAACCGTTCNGAANTTCCNGGAAATTGTNGNANCA
ANAANNNNGAANAAATANANATAAAGG

Blastx:

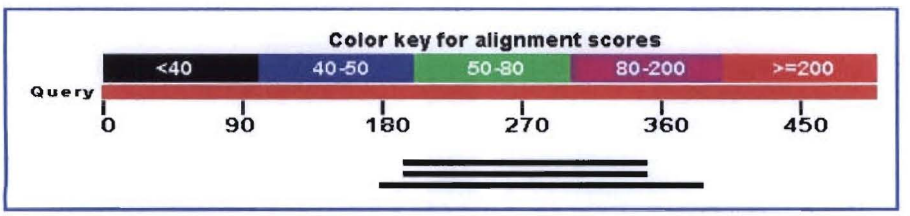
```
Query=
Length=992

No significant similarity found. For reasons why, click here.
```

bsuCST-11:

Sequence (Reverse Complement):

AAGAACATTTCTTAAATCTTAAGCAATAGATTTTTGATTTATTGACAAAGAATAGAAATTCAAAATTCAA
TAATACTCACAATTTAAACAGCATAAACAATATTGAGATATTAAAGAAAAATCATACGCATGAACATAA
CACACAGAAAATTCTCACTACAGCCTATATTGAAGACGAAACTAATACTTGAACAAATTCTTGAGAGTCT
ATTACGCTTTTGTAATTGAGAGTGAAACCAAGCTTTATATTAATTTCATTCACTACTCTGGCATAGCTCTT
GTTCGACTATATCAAAATTGTGAAGTATCATGGATTGAGAATGATGTTATTAACAATCAGATGATGGACT
ATACAGAATCACATAAGATTCAAGAGAAAGTAGAGATTCTAATTTTATGTTTAGACACCATTCTAGATAT
AATCTTTGATAAATTTTGATCTTTAATCTTTAAATTTTTCTATCAATCATTAAAGCATAAAAAAAAAAAAA
AAGG

Blastx:



Sequences producing significant alignments:	Score (Bits)	E Value
gi|21593009|gb|AAM64958.1| unknown [Arabidopsis thaliana]	32.7	5.7
gi|15229382|ref|NP_191871.1| cyclin-dependent protein kinase ...	32.7	5.7
gi|89302250|gb|EAS00238.1| oxidoreductase, short chain dehydr...	32.3	7.5

RNAi Examination

Sequence for the actin gene of *Oxytricha trifallax* was obtained from NCBI based off the submission of Dr. Prescott (U63568 – Oxytricha trifallax B macronuclear actin I gene). The coding region is indicated in blue (bases 205-1332).

```
1    ccccaaaacc ccaaaacccc ataaggatat attttgacat attatgaatg tgacaaatat
61   aatttaagtc gagactgatt aaaagagtcg attgtggcaa aatatacgcc ttcgtaatta
121  tcaaataaga tttattaggg aaagatatgc aatttaatat aataatttat ttcctataaa
181  tctttctatc tactacacgt acatatgtca gaccaacaaa cttgcgttat tgataatggt
241  tcaggagtcg tcaaggccgg tttcgccggt gaggacgctc caagagctgt attcccatca
301  atcgtcggaa gacccaagaa cgtcagcgct ttgatcggtg ttgattccgc ctcagagtac
361  ctcggagatg aagcccaaca aaagagaggg gtcttaaaga tcttctaccc catcgaacat
421  ggtatcgtta aggattggga tgatatggaa aagatctgga accatacctt ctacgttgag
481  ctcagagttc aacccgatga gcacccagtt ctcttgactg aagccccact caacccaaag
541  actaacagag aaaagatgac tcaaatcatg ttcgagacct tcaatgttcc cgctctctac
601  gttgctatcc aagccgttct atctctctat tccgctggta gaacaaccgg tatcgtttgc
661  gatgctggtg atggagttac ccacaccgtc cccatctatg aaggtttctc aatcccacac
721  gccgtatcaa gaatccaact tgccggtaga gacctcacca ccttcttggc caagcttttg
781  accgagagag gatacaactt cacctcatca gctgagatgg aaatcgttag agacattaag
841  gaaaagctct gcttcgttgc tttggactac gagagtgccc ttaagcaatc acacgacagc
901  tcccaattcg agaagaacta cgaacttcca gatggtaagg ttatcaccat tggaaacgag
961  agattcagat gcccagaata tctcttcaag ccactcgaaa tgaacggaag agaactcgac
1021 tccatccaag acttgaccta caaatcaatc caagagtgcg atgtcgacgt tagaagagac
1081 ctctatcaaa acatcatcct ctccggtggt accaccatgt acgaaggtat cggtgagaga
1141 cttctcaagg aaattgagaa cagagctccc aagtcaatca acgttaaagt tattgccagc
1201 ccagacagaa gattcgccgt atggagaggt ggttctaccc ttacttccct ttcaaccttc
1261 gccagcatgt ggatcactaa agaagactac gacgaaaacg gagcaaccat cgttcacaga
1321 aaatgtatct gaggaatcaa atttagtttt agtatatgaa ttataaagtg tgctgagatt
1381 tgccgccctt tgtgactgcg tctcttcaat actgatactt aaccgattgg ctatggacgg
1441 gttgaatgag atctttataa tgtgtgcaca gtcttagtat ttcttaattt aataatttac
1501 aaaaataact tattttgatc gatggggttt tggggttttg gggttttggg gttttgggg
```

Because dsRNA is desired, primers were designed to lie within the highlighted blue region. Two sets of primers were chosen so that the amplicon from the first PCR reaction can act as the template for the nested reaction with primers containing the T7 promoter sequence overhang.

With the aid of the primer designer program FastPCR, the following primer pairs seemed most compatible. Location of these primers along the actin sequence is provided below with the respective color codes.

Compatible combination of pair primers = 517

>F29(153->175) 5'-gtacctcggagatgaagcccaac
Tm=60.0°C Tm(10)=36.4°C CG%=56.5 MW=7042.6 23 bp: PCR efficiency (quality)=116

>R124(1053<=1074) 5'-agtgatccacatgctggcgaag
Tm=60.2°C Tm(10)=41.4°C CG%=54.5 MW=6784.5 22 bp: PCR efficiency (quality)=127

Length of PCR Product= 922
The Optimal Annealing Temperature of PCR = 60.2°C

Compatible combination of pair primers = 3418

>F36(275->295) 5'-agctcagagttcaacccgatg
Tm=57.5°C Tm(10)=36.0°C CG%=52.4 MW=6415.2 21 bp: PCR efficiency (quality)=110

>R112(960<=980) 5'-acgttgattgacttgggagct
Tm=57.1°C Tm(10)=37.1°C CG%=47.6 MW=6492.3 21 bp: PCR efficiency (quality)=93

Length of PCR Product= 706
The Optimal Annealing Temperature of PCR = 59.2°C

```
1     ccccaaaacc ccaaaacccc ataaggatat attttgacat attatgaatg tgacaaatat
61    aatttaagtc gagactgatt aaaagagtcg attgtggcaa aatatacgcc ttcgtaatta
121   tcaaataaga tttattaggg aaagatatgc aatttaatat aataatttat ttcctataaa
181   tctttctatc tactacacgt acatatgtca gaccaacaaa cttgcgttat tgataatggt
241   tcaggagtcg tcaaggccgg tttcgccggt gaggacgctc caagagctgt attcccatca
301   atcgtcggaa gacccaagaa cgtcagcgct ttgatcggtg ttgattccgc ctcagagtac
361   ctcggagatg aagcccaaca aaagagaggg gtcttaaaga tcttctaccc catcgaacat
421   ggtatcgtta aggattggga tgatatggaa aagatctgga accatacctt ctacgttgag
481   ctcagagttc aacccgatga gcacccagtt ctcttgactg aagccccact caacccaaag
541   actaacagag aaaagatgac tcaaatcatg ttcgagacct tcaatgttcc cgctctctac
601   gttgctatcc aagccgttct atctctctat tccgctggta gaacaaccgg tatcgtttgc
661   gatgctggtg atggagttac ccacaccgtc cccatctatg aaggtttctc aatcccacac
721   gccgtatcaa gaatccaact tgccggtaga gacctcacca ccttcttggc caagcttttg
781   accgagagag gatacaactt cacctcatca gctgagatgg aaatcgttag agacattaag
841   gaaaagctct gcttcgttgc tttggactac gagagtgccc ttaagcaatc acacgacagc
901   tcccaattcg agaagaacta cgaacttcca gatggtaagg ttatcaccat tggaaacgag
961   agattcagat gcccagaata tctcttcaag ccactcgaaa tgaacggaag agaactcgac
1021  tccatccaag acttgaccta caaatcaatc caagagtgcg atgtcgacgt tagaagagac
1081  ctctatcaaa acatcatcct ctccggtggt accaccatgt acgaaggtat cggtgagaga
1141  cttctcaagg aaattgagaa cagagctccc aagtcaatca acgttaaagt tattgccagc
1201  ccagacagaa gattcgccgt atggagaggt ggttctaccc ttacttccct ttcaaccttc
1261  gccagcatgt ggatcactaa agaagactac gacgaaaacg gagcaaccat cgttcacaga
1321  aaatgtatct gaggaatcaa atttagtttt agtatatgaa ttataaagtg tgctgagatt
1381  tgccgccctt tgtgactgcg tctcttcaat actgatactt aaccgattgg ctatggacgg
1441  gttgaatgag atctttataa tgtgtgcaca gtcttagtat ttcttaattt aataatttac
1501  aaaaataact tattttgatc gatggggttt tggggttttg gggttttggg gttttgggg
```

Through the use of Integrated DNA Technologies' OligoAnalyzer, further evaluation of these primer choices can be done.

Primer Set One:
Forward primer: T_m = 59.4°C; 4 possible hairpin structures; 13 possible self-dimers
Reverse primer: T_m = 59.8°C; 3 possible hairpin structures; 14 possible self-dimers
Heterodimers: 15 possible heterodimer structures

Primer Set Two:
Forward primer: T_m = 57.0°C; 5 possible hairpin structures; 11 possible self-dimers
Reverse primer: T_m = 56.7°C; 1 possible hairpin structure; 4 possible self-dimers
Heterodimers: 13 possible heterodimer structures

The sequence for the T7 promoter necessary for future construction of dsRNA is 'gaattaatac gactcactata gggaga' (15), so was added to both primers of the second set. The following primers were ordered from Integrated DNA Technologies.

1) ActinF1: 5'- gtacctcgga gatgaagccc aac
2) ActinR1: 5'- agtgatccac atgctggcga ag
3) ActinF2: 5'- agctcagagt tcaacccgat g
4) ActinF2T7: 5'- gaattaatac gactcactat agggagaagc tcagagttca acccgatg
5) ActinR2: 5'- acgttgattg acttgggagc t
6) ActinR2T7: 5'- gaattaatac gactcactat agggagaacg ttgattgact tgggagct

50 µl PCR reactions were set up with the first primer set (F1 and R1) and second primer set (F2 and R2), varying conditions from those that had been optimized by Prescott (14) and those optimized by others (15). Stock JRB 310 DNA was used to make serial dilutions to run multiple PCR reactions for optimization. The final concentration of dNTPs used varied from 10 to 50 mM, the final concentration of $MgCl_2$ varied between 1 and 15mM, and primer concentration between 1 ug and 0.1ng. A range of annealing temperatures (48°C to 60°C) was also run in a hope to optimize the PCR reaction.

After PCR was completed, 5 µl of the reaction was combined with 4 µl of TE buffer and 1 µl of Blue Juice and run at 75 Volts for 40 minutes on a 2% agarose gel (with 0.5 µg/ml ethidium bromide) alongside 8 µl of a 1 Kb ladder.

Unfortunately, after all the attempts to optimize the reaction, there was still no amplicon product visible on the gel alongside the ladder.

Discussion

cDNA Library Analysis

There were several difficulties while performing the research for this thesis. Feelings were optimistic that this first run through with the Superscript™ Plasmid System for cDNA synthesis and cloning were successful after screening on agar plates for transformation. Not only were the initial plates to screen for successful transformation spotted with colonies, but also after expansion, the plates became lawns of bacterial growth. However, this only means that the DH5α cells took up the pCMV•SPORT 6 vector that contains ampicillin resistance. It does not look at whether the cDNA created along the process of the protocol was inserted into the vector.

Invitrogen designed this specific cloning vector to provide a low background of nonrecombinant colonies. However, with our sequencing, 10 out of 12 of the randomly selected transformed colonies contained no cDNA insert. Even more puzzling is that among those samples that did not contain a cDNA insert, they did contain a common sequence of 19 base pairs (TCGCGATCTATAACTATTC). It is unknown if this sequence is part of the vector and is not reported in the manual as part of the vector sequence, or whether there are junk fragments from the adapter steps that were incorporated into the vector during the ligation step. The column chromatography steps were supposed to prevent shorter fragments (unligated adapters, small cDNAs, and the primer-adapter fragment released from the cDNA by Not I digestion) from becoming incorporated. It may prove worthy to look at the early collection tubes from the column chromatography just in case the tubes used in this project had collections too late and large cDNA inserts were lost.

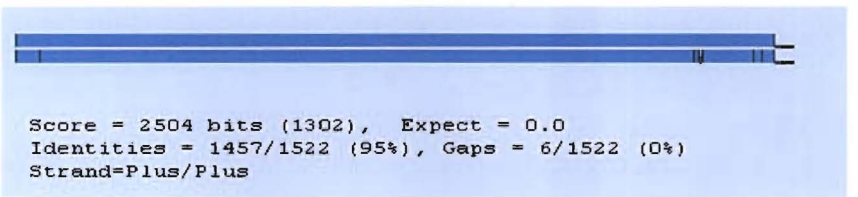
A 0.0004% transformation efficiency and a 2 for 12 success rate in obtaining vector sequence with a cDNA insert may not sound outstanding, but there is still that one sample (bsuCST-11) that poses some interesting questions. Further inquiries will have to be put in to Dr. Orias and others in the group that submitted the theoretical protein sequence for *Tetrahymena*, to provide insight as to a function for our cDNA that has such a similarity.

As we continue to perfect the protocol for creating these cDNA libraries, in the future the libraries will expand to different developmental stages of *Oxytricha* and more will be learned about the genetic mechanisms surrounding development and intracellular patterning and organization.

RNAi Examination

Like the analysis of the cyst cDNA library, there were just as many problems with this aspect of the thesis. Unfortunately no PCR product could be obtained. The first possibility is that there was a problem with the DNA. If the DNA was suspended in EDTA, this could present problems for $MgCl_2$ in the PCR reaction.

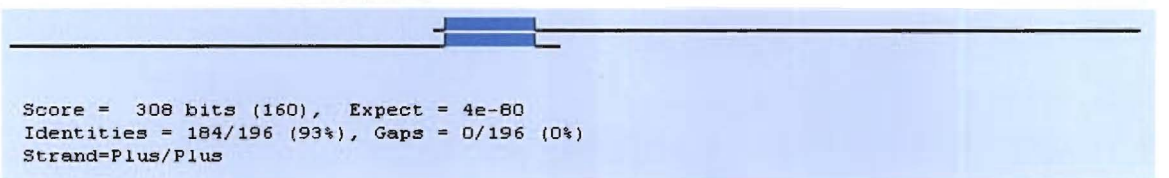
Another possibility is that primers were not designed to the correct sequence for the DNA at hand, JRB 310. There is variability in the sequences submitted to NCBI for the actin gene by species. An alignment was performed on *Oxytricha trifallax* B macronuclear actin I gene (U63568), which was used for designing these primers, and *Oxytricha fallax* macronuclear actin I gene (U63567).



```
Score = 2504 bits (1302),  Expect = 0.0
Identities = 1457/1522 (95%), Gaps = 6/1522 (0%)
Strand=Plus/Plus
```

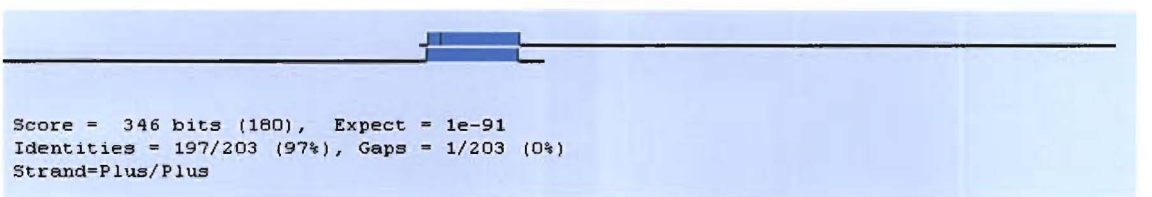
If variations in sequence from line to line allows for a large enough gap that the primer set was trying to anneal to, no product will be possible. To appreciate the possibilities of variation in sequence between populations and between the macronuclear actin gene and micronuclear actin gene, further alignments were performed on *Oxytricha fallax* micronuclear actin I gene (U63581)

Alignment of U63568 and U63581:



```
Score =  308 bits (160),  Expect = 4e-80
Identities = 184/196 (93%), Gaps = 0/196 (0%)
Strand=Plus/Plus
```

Alignment of U63567 and U63581:



```
Score =  346 bits (180),  Expect = 1e-91
Identities = 197/203 (97%), Gaps = 1/203 (0%)
Strand=Plus/Plus
```

Further PCR will need to be performed to determine whether there is a DNA or primer problem. However, once the problem is fixed, dsRNA can be produced and experiments run on individual ciliates to determine whether it is possible to induce RNA interference by direct exposure to dsRNA (without a need for introducing E. coli to *Oxytricha* as a food source). If this is possible, then there is a means to pull out genes from cDNA libraries at each developmental stage and determine function by phenotypic observation by targeting the gene through RNAi.

References

1. Grimes, G.W. 1972. Cortical structure in nondividing and cortical morphogenesis in dividing *Oxytricha fallax*. *J Protozool* 19:428-445.
2. Sonneborn, T.M. 1964. The differentiation of cells. *Proc Natn Acad Sci* 51:915-929.
3. Grimes, G.W. 1973. Differentiation during encystment and excystment in *Oxytricha fallax*. *J Protozool* 20:92-104.
4. Villalobo, E., L. Morin, C. Moch, R. Lescasse, M. Hanna, W. Xiao and A. Baroin-Tourancheau. 2002. A homologue of CROC-1 in a ciliated protist testifies to the ancient origin of the Ubiquitin-conjugating enzyme variant family. *Mol Biol Evol* 19:39-48.
5. Hartshorn, C., A. Anshelevich and L.J. Wangh. 2005. Rapid, single-tube method for quantitative preparation and analysis of RNA and DNA in samples as small as one cell. *BMC Biotech* 5:1-13.
6. Ammermann, D. 1971. Morphology and development of the macronuclei of the ciliates *Stylonychia mytilus* and *Euplotes aediculatus*. *Chromosoma* 33:209-238.
7. Chalker, D.L., P. Fuller, M.C. Yao. 2005. Communication between parental and developing genomes during *Tetrahymena* nuclear differentiation is likely mediated by homologous RNAs. *Genetics* 169:149-160.
8. Klobutcher, L.A. and G. Herrick. 1995. Consensus inverted terminal repeat sequence of *Paramecium* IESs: resemblance to termini of Tc1-related and *Euplotes* Tec transposons. *Nucleic Acids Research* 23:2006-2013.
9. Couzin, J. 2002. Small RNAs make big splash. *Science* 298:2296-7.
10. Galvani A and L. Sperling. 2002. RNA interference by feeding in *Paramecium*. *TRENDS in Genetics* 18:11-12.
11. Paschka, A.G., F. Jonsson, V. Maier, M. Mollenbeck, K. Paeschke, J. Postberg, S. Rupprecht and H.J. Lipps. 2003. The use of RNAi to analyze gene function in spirotrichous ciliates. *Europ J Protistol* 39:449-454.
12. Villalobo, E., C. Moch, R. Perasso and A. Baroin-Tourancheau. 2001. Searching for excystment-regulated genes in *Sterkiella histriomuscorum* (*Ciliophora, Oxytrichidae*): a mRNA differential display analysis of gene expression in excysting cells. *J. Eukaryot. Microbiol.* 48:382-390.
13. Roskams, J and L. Rodgers. Lab Ref: A handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench. New York: Cold Spring Harbor Press, 2002.
14. DuBois, M.L. and D.M. Prescott. 1997. Volatility of internal eliminated segments in germ line genes of Hypotrichous ciliates. *Molecular and Cellular Biology* 17:326-337.
15. Clemens, J.C., C.A. Worby, N. Simonson-Leff, M. Muda, T. Maehama, B.A. Hemmings and J.E. Dixon. 2000. Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *PNAS* 97:6499-6503.